Switching cell penetrating and CXCR4-binding activities of nanoscale-organized arginine-rich peptides

Marianna Teixeira de Pinho Favaro [1, 2 π], Naroa Serna [2, 3, 4 π], Laura Sánchez-García [2, 3, 4], Rafael Cubarsi [5], Mónica Roldán [6¥], Alejandro Sánchez-Chardi [6], Ugutz Unzueta [4, 7], Ramón Mangues [4, 7], Neus Ferrer-Miralles [2, 3, 4], Adriano Rodrigues Azzoni [8], Esther Vázquez [2, 3, 4 *], Antonio Villaverde [2, 3, 4 *]

[1] Centro de Biologia Molecular e Engenharia Genética, Universidade Estadual de Campinas, Av Candido Rondon, 400, 13083-875 Campinas, SP, Brazil.

[2] Institut de Biotecnologia i de Biomedicina (IBB), Universitat Autònoma de Barcelona, 08193 Cerdanyola del Vallès, Spain.

[3] Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, 08193 Cerdanyola del Vallès, Spain.

[4] CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), 08193 Cerdanyola del Vallès, Spain.

[5] Departament de Matemàtiques, Campus Nord C3-212, Universitat Politècnica de Catalunya, 08034 Barcelona, Spain.

[6] Servei de Microscòpia, Universitat Autònoma de Barcelona, 08193 Cerdanyola del Vallès, Spain.

[7] Institut d'Investigacions Biomèdiques Sant Pau and Josep Carreras Research Institute, Hospital de la Santa Creu i Sant Pau, 08025 Barcelona, Spain.

[8] Departamento de Engenharia Química, Escola Politécnica, Universidade de São Paulo, Av. Prof. Luciano Gualberto, Trav. 3, Nº 380, 05508-900, São Paulo, SP, Brazil.

[π] Equally contributed

[¥] Present address: Hospital Sant Joan de Déu, Passeig de Sant Joan de Déu, 2, 08950 Esplugues de Llobregat, Barcelona

[*] Corresponding authors:
EV, Esther.Vazquez@uab.es; AV, Antoni.Villaverde@uab.es

WORD COUNT (to include body text and figure legends),
ABSTRACT: 146
TEXT: 4633
REFERENCES: 42
FIGURES: 6

The authors do not appreciate any conflict of interest.

ABSTRACT

Arginine-rich protein motifs have been described as potent cell-penetrating peptides (CPPs) but also as rather specific ligands of the cell surface chemokine receptor CXCR4, involved in the infection by the human immunodeficiency virus (HIV). Polyarginines are commonly used to functionalize nanoscale vehicles for gene therapy and drug delivery, aimed to enhance cell penetrability of the therapeutic cargo. However, under which conditions these peptides do act as either unspecific or specific ligands is unknown. We have here explored the cell penetrability of differently charged polyarginines in two alternative presentations, namely as unassembled fusion proteins or assembled in multimeric protein nanoparticles. By this, we have observed that arginine-rich peptides switch between receptor-mediated and receptor-independent mechanisms of cell penetration. The relative weight of these activities is determined by the electrostatic charge of the construct and the oligomerization status of the nanoscale material, both regulatable by conventional protein engineering approaches.

Keywords: Protein materials; protein engineering; self-assembling; CXCR4; tumour-homing peptides

INTRODUCTION

The cell surface chemokine receptor CXCR4 is of high clinical relevance, as it acts as a co-receptor during the cell infection by the human immunodeficiency virus (HIV) [1]. In addition, CXCR4 is a pivotal cancer marker overexpressed in the stem cells of more than 20 human neoplasias, including frequent ones such as colorectal and pancreatic cancers and lymphoma [2]. In colorectal cancer, overexpression of CXCR4 is correlated with aggressiveness and metastatic potential. This fact has attracted interest over this membrane protein as a potential target for drug delivery [3-5]. Polyarginines (like R9 and others) are cell penetrating peptides (CPPs) that being cationic, efficiently condensate nucleic acids and promote nuclear localization of attached molecules [6]. Such pleiotropic profiling has empowered some members of the Rn series as valuable functionalizing agents in gene therapy and non-targeted drug delivery. However, under some experimental conditions, polyarginines such as R12 (but not R8), presented as free synthetic peptides, are endocyted by cultured cells in a CXCR4-dependent fashion [7]. In addition, it has been recently reported that a fraction of R9-displaying bacterial amyloids internalize cultured cells by direct binding to CXCR4 [8], while a part of the material penetrates by rather unspecific mechanisms. The possibility to specifically mediate the internalization of macromolecular complexes via CXCR4 is highly appealing in the context of drug delivery. Then, appropriate polyarginine candidates might represent novel and valuable peptidic ligands for CXCR4 that might offer innovative pharmacological opportunities over the currently explored peptide ligands of CXCR4, whose performance is highly variable [9]. In this context, it would be very convenient to define the optimal form and presentation of a given polyarginine domain to unbalance the alternative routes of cell penetration to favour the specific, CXCR4-dependent cell penetrability. This is of special interest when polyarginines are linked to high molecular weight cargos such as full-length recombinant proteins or different types of nanoparticles or other nanostructured entities. Then, we have explored here

the dependence on CXCR4 in the internalization of four members of the Rn family when fused to a soluble functional GFP. The fluorescent protein acts as a convenient reporter for image-based monitoring and analysis and as an efficient building block for the generation of protein-based nanoparticles [10]. We have specially focused our study on the length of the cationic segment (the n value) and on the multivalent versus monovalent display of the ligand on the surface of the protein material. By taking this approach, we have identified both parameters as determinants of the CXCR4 specificity in the cellular penetrability of Rn-empowered constructs.

METHODS

Proteins and protein production

Four GFP-derived fusion proteins, namely R3-GFP-H6, R6-GFP-H6, R7-GFP-H6 and R9-GFP-H6 [11] were used in the present study upon recombinant production in bacteria. As previously described, R9-GFP-H6 was modified by directed mutagenesis to generate the other constructions, replacing arginines by glycines and alanines, in order to maintain the same length of peptide tag with different charges [11]. All of the fusion proteins are based on the same modular scheme (Figure 1A), in which the cationic peptide is placed at the N-terminal end of a His-tagged GFP. For highly cationic peptide segments such arrangement promotes the self-assembling of GFP as highly stable protein nanoparticles, which act as multivalent materials usable as vehicles for cell-targeted drug delivery [12-14]. *Escherichia coli* Rosetta (Novagen, Madison, WI, USA) was grown in shaker flask in Lysogenic broth (LB) medium containing 34 mg/ml chloramphenicol, 12.5 mg/ml tetracycline (strain resistance) and 100 mg/ml ampicillin (vector resistance) at 37 ºC 250 rpm, to reach an OD_{550} of ~0.6 units. Then, induction of gene expression was triggered by 0.1 mM isopropyl-b-D-thiogalactopyronaside (IPTG, Merck, Kenilworth, NJ, USA), and it was prolonged overnight at 20 ºC. Bacterial cells were collected by centrifugation and resuspended in Wash buffer (20 mM Tris , 500 mM NaCl, 10 mM Imidazole; pH 8.0) in the presence of EDTA-Free protease inhibitor (Complete EDTA-Free; Roche, Manheim, Germany), to be disrupted at 1200 psi using a French Press (Thermo FA-078A, Thermo Electron Corporation, Needham Heights, MA, USA). The soluble fraction of lysed cells was loaded on a HiTrap Chelating HP 1 ml column (GE Healthcare, Chicago, IL, USA) to be purified by 6x His-tag affinity chromatography on an ÄKTA purifier (GE Healthcare, Chicago, IL, USA). Bound proteins were eluted with Elution Buffer (20 mM Tris , 500 mM NaCl, 500 mM Imidazole; pH 8.0) in a linear gradient, and fractions containing the protein were then dialysed overnight at 4 ºC against Tris

Dextrose buffer (20 mM Tris 5 % dextrose; pH 7.4, referred as *Dextrose* in the Figures) or Tris NaCl buffer (20 mM Tris, 500 mM NaCl; pH 7.4, referred as *NaCl* in the Figures), depending on the solubility of each protein. The purity of the protein was determined by denaturing SDS-polyacrylamide gel electrophoresis (12 % polyacrylamide) and anti-6x-His-tag Western Blot. Concentrations were determined by Bradford's assay. R9-GFP-H6 had been extensively characterized regarding several of its physicochemical properties, including self-assembling [15].

Fluorescence determination and dynamic light scattering (DLS)

The specific fluorescence of protein variants was determined in a Varian Cary Eclipse Fluorescence Spectrophotometer (Agilent Technologies, Santa Clara, CA, USA) with all proteins being diluted to the same concentration (0.1 mg/ml). For measurements, samples were excited at a wavelength of 488 nm and the emission collected in the range 500-548 nm, with maximum emission detected at 510 nm. Volume size distribution of nanoparticles and monomeric proteins were determined by dynamic light scattering at 633 nm (Zetasizer Nano ZS, Malvern Instruments Limited, Malvern, Worcestershire, UK). Measurements were performed in triplicate. For disassembling assays, SDS (to 0.1 % or 0.05 %, final) was added to proteins previously diluted at 1 mg/ml, and the samples were incubated for 10 min and finally submitted to light scattering analysis.

Ultrastructural characterization

Morphometry (size and shape) of representative nanoparticles was evaluated quantitatively and qualitatively both with field emission scanning electron microscopy (FESEM) and transmission electron microscopy (TEM). Drops of 3 l of R7-GFP-H6 in both Tris Dextrose and Tris NaCl buffers and R9-GFP-H6 in Tris Dextrose buffer were directly deposited on silicon wafers (Ted Pella Inc., Redding, CA, USA) for 1 min, excess blotted with Whatman filter paper number 1 (GE Healthcare, Little

Chalfont, UK), air dried, and immediately observed without coating at a nearly native state in a FESEM Zeiss Merlin (Zeiss, Oberkochen, Germany) operating at 1 kV with a high resolution in-lens secondary electron detector. Drops of 3 l of the same three samples were directly deposited on 200 mesh carbon-coated copper grids (Electron Microscopy Sciences, Hatfield, PA, USA) for 1 min, excess blotted with Whatman filter paper number 1 (GE Healthcare), contrasted with 3 l of 1 % uranyl acetate (Polysciences Inc., Warrington, PA, USA) for 1 min, blotted again and observed in a TEM Jeol 1400 (Jeol Ltd., Tokyo, Japan) operating at 80 kV with a Gatan Orius SC200 CCD camera (Gatan Inc. Abingdon, UK), For each sample and EM technique, the size of 50 randomly distributed particles was measured from TIFF and DM3 files with ImageJ software (NIH, Bethesda, MA,USA).

Cell culture and internalization

HeLa (ATCC-CCL-2, ATCC, Manassas, VA, USA) cells were cultured in MEM-alpha (GIBCO) supplemented with 10 % Foetal Calf Serum (GIBCO BRL, Grand Island, NY, USA) and incubated at 37ºC and 5 % CO_2. For internalization assays, cells were grown in 24-well plates on complete medium, which was replaced by OptiPro supplemented with L-Glutamine before the addition of the proteins. After 2 or 24 h of incubation, cells were detached from the plate with Trypsin (GIBCO) 1 mg/ml for 15 min before being analysed by flow cytometry on a FACS Canto (Becton Dickinson, Franklin Lakes, NJ, USA). Protein fluorescence was excited using a 15 mW air-cooled argon ion laser at 488 nm and detected by a 530/30 nm band pass filter D detector. In all internalization experiments the same concentration of protein was used (1 or 2 M, as specified). The results were corrected with the fluorescence values obtained from fluorimeter to render data comparable in terms of protein mass. For competition assays, the specific CXCR4 antagonist AMD3100 [16] (octahydrochloride hydrate, Sigma-Aldrich) was added to fresh OptiPRO medium to

a 1:10 (protein: AMD3100) molar ratio. Nanoparticles were added after a 1 h incubation with AMD3100.

Confocal microscopy

For confocal microscopy HeLa cells were grown on Mat-Tek plates (MatTek Corporation, Ashland, MA, USA). The nuclei were labelled with 10 g/ml Hoechst 33342 (Invitrogen, Waltham, MA, USA) and the plasma membrane with 2.5 g/ml CellMask™ Deep Red (Molecular Probes, Eugene, OR, USA) for 10 min at room temperature and then washed in PBS buffer (Sigma-Aldrich Chemie GmbH, Taufkirchen, Germany). Live cells were recorded by TCS-SP5 confocal laser microscopy (Leica Microsystems, Wetzlar, Germany) using a Plan Apo 63x/1.4 (oil HC x PL APO lambda blue) objective. Hoechst 33342 DNA label was excited with a blue diode (405 nm) and detected in the 415–460 nm range. GFP-proteins were excited with an Ar laser (488 nm) and detected in the 525–545 nm range. CellMask™ Deep Red was excited with a HeNe laser (633 nm) and detected in the 650–775 nm range. To determine the protein localization inside the cell, stacks of 20–30 sections were collected with 0.5 µm of thickness, and 3D models were generated using the Leica LAS X software (Leica Microsystems, Wetzlar, Germany).

Statistical analysis

Values are expressed as mean data and standard error ($\bar{x} \pm SD$). Data were tested for normal distribution and for homogeneity of variance prior the use of parametric tests with Kolmogorov-Smirnov and Levene tests, respectively. Multiple comparisons were performed by one-way ANOVAs followed by Fisher's least significant difference (LSD, two tailed) and pairwise comparisons by two-tailed Student t-tests using Microsoft Excel 2011 and SPSS 15.0 softwares.

Numerical modelling

To model the penetration of protein nanoparticles in a kinetic way we selected the simplest model by assuming a constant rate of nanoparticle uptake $m(t)$ as a function of time t,

$$\frac{dm}{dt} = -k, 0 \leq k < 1 \Rightarrow m(t) = Me^{-kt} \quad (1)$$

The intracellular fluorescence emission $f(t)$ is proportional to the amount of internalized protein,

$$f(t) = L\,(1 - e^{-kt}) \quad (2)$$

and the initial entry rate of fluorescent materials is

$$\left(\frac{df(t)}{dt}\right)_{t=0} = Lk \equiv v \quad (3)$$

The value of k refers to the time $\tau = ln2/k$ needed to reach one half of the maximum fluorescence L accumulated in target cells. The model depends on the parameter k and the integration constant L. To determine k and L, two couples of measurements (2 h and 24 h) were employed, namely t_1, $f_1 = f(t_1)$, and t_2, $f_2 = f(t_2)$. We assume $t_1 < t_2$; $f_1 < f_2$. Therefore, according to Eq. (2),

$$k = -\frac{1}{t1}\ln\left(1 - \frac{f1}{L}\right) = -\frac{1}{t2}\ln\left(1 - \frac{f2}{L}\right) \quad (4)$$

By defining $x = f_1/L < 1$, $\mu = t_2/t_1 > 1$, and $\alpha = f_2/f_1 > 1$, the second equality in the above equation leads to the following non-linear equation

$$(1 - x)^{\mu} + \alpha x - 1 = 0 \quad (5)$$

If μ is an integer, the above equation is a polynomial. In a general case, a fixed point Newton-Raphson method will be applied to obtain the single root x_0 satisfying $0 < x_0 < 1$. Finally, we get $k = -\frac{1}{t1} ln(1 - x_0)$, $L = \frac{f1}{x0}$.

RESULTS

R3-GFP-H6, R6-GFP-H6, R7-GFP-H6 and R9-GFP-H6 (Figure 1, A) were successfully bio-produced in *E. coli* (Figure 1B) and stored in either Tris Dextrose buffer (named as *Dextrose* in the Figures) or Tris NaCl buffer (named as *NaCl* in the Figures), depending on their solubility. R3 and R6 derivatives were preferentially soluble in Tris NaCl buffer, R9 in Tris Dextrose buffer and R7 was found to be soluble in both (data not shown). All produced proteins were fluorescent, although with important variability between the checked species (Figure 1, C). This observation, strongly suggested a differential impact (direct or indirect) of the differently charged amino terminal Rn tails on the folding status of the GFP variants. Since the buffer is not expected to affect fluorescence by itself, oligomerization and associated conformational changes might affect the emission capacity of the fluorophore or quench it. When analysing by DLS the potential of these constructs to form nanoparticles, linked to the nature of the cationic terminal domain, we confirmed the inability of R3 and R6 to promote self-assembling (Figure 2), in agreement with a previous preliminary screening [11]. Interestingly, R7-GFP-H6 showed self-assembling properties in Tris Dextrose resulting in nanoparticles of about 30 nm, but not in Tris NaCl in which the protein remained unassembled (Figure 2, and Figure 3, A and B). The higher salt content in the latest buffer might interfere with the electrostatic interactions needed between building blocks to start the oligomerization process [11], thus preventing nanoparticle formation. As previously described [15], R9-GFP-H6 efficiently assembled as regular, toroid-shaped materials of about 30 nm (Figure 2). The size measurements of nanoparticles obtained from FESEM and TEM images were in deep agreement with the values obtained by DLS (Figures 2, 3). In Dextrose-containing buffer, the R9-GFP-H6 nanomaterial appeared larger than in conventional buffers [15], as the sugar probably stabilizes protein-protein cross-interactions. Interestingly, the protein versions that self-organize as oligomers rendered lower fluorescence emission levels than those that remained as building

blocks (Figure 1, C). This fact indicated that the architectonic organization of the whole material has an impact on the fluorescent emission of the core GFP, either by a perturbation on the individual building block conformation, as suggested by previous analyses of related self-assembling fluorescent proteins [10, 17], or by fluorophore quenching due to specific protein-protein cross interactions [18]. Irrespective of these differences, the fluorescence emission was, in all cases, high enough to monitor fast and accurately the cell internalization process upon exposure, and for the evaluation of the potential mediation of CXCR4 in the uptake process.

Cell penetrability of all these constructs was explored by the accumulation of intracellular fluorescence in cultured HeLa cells exposed to them, after a harsh trypsin treatment to remove externally associated protein. As observed (Figure 4A), all proteins penetrate cells since early contact times in a dose-dependent manner, with efficacies that clearly depend on the number of arginines composing the tag at the protein amino terminus. Previous studies showed that 6-12 arginine residues are necessary to promote unspecific entry, with guanidino groups forming hydrogen bonds and ion-pairs with lipid head groups and hydrophobic tails to assist in membrane binding [19]. Here we observed that differences between proteins were largely amplified at long exposure times (24 h), especially in the case of R9-GFP-H6, that was clearly superior than the related polypeptides in the uptake.

Although the number of arginine residues had been implicated in membrane permeability [20], there was a lack of systematic comparisons to clearly understand the role of oligomerization in promoting internalization and how the entry mechanism is affected by these factors. At 24 h, a penetrability of the assembled R7-GFP-H6 (in Tris Dextrose) higher than the unassembled version (in Tris NaCl) was evidenced, pointing out the assembled form and the multivalent display of R7 as favouring agents of cell uptake. Although the number of arginine residues was a parameter that

positively affected internalization, multimerization appeared to have an additive and more potent weight. Confocal imaging of the penetration process (Figure 4 B and C) revealed clustering of fluorescent materials in the perinuclear region essentially in the case of the R9-based construct, with an intracellular distribution that might be compatible with both endosomal uptake and transmembrane penetration. In this regard, it must be stressed that endocytic vesicles engulfing fluorescent material were hardly identified during the penetration of R9-GFP-H6, while they were more apparent in the penetration of the assembled R7-GFP-H6. In this regard, it must be noted the merging yellow signal (co-localization between red membrane and green protein, arrows) revealed in confocal images of cells exposed to R7-GFP-H6 nanoparticles (Figure 4 B).

It has been reported that CXCR4 is involved in internalization of R12 through macropinocytosis, while R8 and TAT seemed to follow a CXCR4-independent route [7]. To determine the involvement of CXCR4 as mediator of internalization of R7 and R9, we monitored this event in presence of AMD3100, a chemical ligand of CXCR4 that inhibits the binding of CXCR4-reactive proteins [16, 21]. As observed (Figure 5, A), at early times after exposure the CXCR4 specificity of the uptake of monomeric versions increases with the number or arginine residues in a linear way. However, the multivalent presentation of polyarginines seemed to promote receptor-independent penetrability, probably linked to the CPP nature of Rn, since the intracellular accumulation of fluorescence is poorly prevented by AMD3100 in the case of the oligomers. However, after 24 h of exposure, the inhibitory effect of AMD3100 was clearly more evident in the case of R9-GFP-H6 (and also R7-GFP-H6) than in the rest of the materials. In this context, it must be noted that the assembled R7 and R9 versions of GFP are indistinguishable regarding stability, as both are equally dissociated by different concentrations of SDS (Figure 5, B). Therefore, dissociation for R7 and R9 constructs (and differential dissociation) was

not expected under these experimental conditions. Since both the number of N-terminal arginine residues and the oligomerization process might influence efficiency and specificity of protein penetration is highly relevant for the design of cell-targeted nanoparticle, we presented the obtained data (from Figure 5, A) in a more visual way for further analysis.

As observed (Figure 6, A), the number of arginine residues positively influenced the amount of internalized fluorescence both 2 h and 24 h upon exposure. Oligomerization, affecting only R7- and R9-based materials, showed a moderate impact on the global uptake process that was dissimilar when comparing short and long times. On the other hand, the % of uptake inhibition mediated by AMD3100 also increased along the number of arginine residues, at 24 h but not from the determinations done 2 h upon exposure. While the specificity for CXCR4 appears as being globally gained by the accumulation of cationic residues, that progressively convert Rn-based CPPs into CXCR4 ligands, data also suggested a differential uptake mechanisms acting at short and long incubation times. To better analyze this possibility, we modelled the penetration of protein nanoparticles in a kinetic way. For that, we explored the factors *L, k, τ* and *v* as defined in the materials and methods section. When these parameters were determined versus the number of arginine residues at the N terminus of the proteins (Rn) for exposures in absence and in presence of the inhibitor AMD3100, we obtained graphical determinations of their behaviour (Figure 6, B). Importantly, the global amount of internalized protein *L* increased with the Rn value, confirming the positive impact that the number of arginines has in the global penetrability of proteins, that appears as being further (but slightly) enhanced by the oligomerization of the building blocks.

Irrespective of the precise mechanism of entrance, this can be accounted by the increase in the positive charge, but also by the multivalent exposure of Rn versus a

monovalent display in unassembled proteins [3]. On the other hand, the initial penetration velocity of proteins (v) increased with Rn values for unassembled proteins but it was inversely proportional to the number of arginines in the nanoparticles. In presence of AMD3100, that blocks CXCR4-specific penetration and only unspecific CPP-based uptake is allowed, v tends to be constant. This is indicative that the number or arginines impacts only (or majorly) on the receptor-dependent penetration of the proteins, which starts very fast upon protein-cell contact. In addition, the comparative behaviour of v in absence and in presence of AMD3100 also indicates that oligomerization globally enhances the CPP properties of proteins, a fact that promotes a fast and early entrance into the cells upon exposure. Inversely, oligomerization, and the increase of the number of arginine residues in the oligomers minimized specificity in early penetration stages. By analyzing k, which can be also interpreted in terms of the half time (τ) to reach the maximal intracellular accumulation of protein, we determined that the jump from R7 to R9 (but not oligomerization itself) expands the time period in which the penetration process does occur. However, when AMD3100 is present, the duration of the penetration process is also expanded depending on the Rn value. Therefore, the unspecific penetrability of Rn-based nanoconstructs sustained by their CPP properties is a time-prolonged process that is favoured when the CXCR4-dependent endosomal penetration is not available.

DISCUSSION

The delivery of therapeutic molecules into cells requires the smart engineering of fusogenic agents, mainly lipids [22, 23] and proteins [24, 25], that act as unspecific but highly efficient cell penetrating agents. Alternatively, cell-targeting tools, such as antibodies or peptidic ligands, provide selectivity in the cell binding of drug vehicles [26, 27], although they are generically less competent than unspecific CPP tools in promoting internalization [28, 29]. The still poorly explored combination of both agents in the same vehicle, namely the incorporation of a CPP and a peptidic ligand of a target cell surface receptor has so far resulted in very high cell penetration levels but at expenses of specificity [30].

Among the CPPs recognized as useful and with potential for realistic development and applicability in biomedicine, arginine-rich peptides are indeed being translated from bench to bedside in several pre-clinical and clinical trials [31]. Even though the internalization mechanisms of polyarginines are not completely understood, they are being combined to different formulations due to their penetration enhancing properties, being applied even in highly challenging conditions such as oral administration [32]. Interestingly, it has been separately suggested that polyarginines exhibit either unspecific (CPP) or specific (receptor-dependent) cell penetration activities [7, 8]. Receptor-dependent internalization of polyarginines can be seen as an entangled scenario affected by multiple factors. For instance, the contribution of Syndecan-4, a receptor known to be involved in R8 internalization, was also reported to depend on parameters such as the extracellular concentration of the peptide but not affected by the presence of a protein fused to R8 [33].

This unusual pleiotropic profile might be highly relevant when designing new generation protein-based vehicles for cell-targeted drug delivery, especially regarding efficiency and when looking for selectivity in a drug delivery process. Therefore, the

highly versatile mechanism of entry of polyarginines has been dissected here through the use of several Rn constructs, showing or not self-assembling properties that affect the multivalent display of the peptide to exposed cells. In this context, the obtained data indicate the coexistence of two different mechanism of penetrability of Rn-empowered proteins, namely CXCR4-dependent uptake and unspecific CPP-based internalization, which act differentially during the time of contact between cells and materials. In presence of free CXCR4 on the cell surface, proteins probably remain attached to the receptor and initiate a fast process of penetration. When CXCR4 is blocked by AMD3100, or CXCR4 paratopes are saturated by earlier contacts, cell penetrability is unspecific and takes place much slower (Figure 5). The relative prevalence of both mechanisms is also largely influenced by the value of Rn and by the oligomerization status of the protein. The specificity in the cell penetration is reduced by oligomerization and by the Rn value of the oligomers at short times upon exposure (Figure 6). However, globally, the increase of Rn length and the formation of multivalent structures increase the total amount of intracellular material and its receptor-depended uptake.

In summary, despite polyarginines have been generically observed as highly potent CPPs useful as an internalization tag in cell therapy and drug delivery [6], their known residual specificity for the cell surface cytokine receptor CXCR4 can be enhanced by a proper presentation of the cationic stretch. As shown here, this can be achieved by extending the number of arginine residues and, with a milder impact, of the multimeric presentation on the surface of targeting vehicles. Since polyarginines are highly cationic and they act also as oligomerization domains [9], R9 and related species can confer to fusion proteins both self-assembling and cell-targeting properties. Being CXCR4 an appealing target in innovative cancer therapies and in antiretroviral treatments [2, 4, 5, 7, 16, 21, 34-37], R9 and related peptides might represent an additional instrument for the design of improved vehicles

for intracellular drug delivery. Albeit the CXCR4 specificity of polyarginines might not be absolute, even in their optimal presentation (Figure 5, A), their use in combination with other CXCR4 ligands might allow the generation of new-generation bi-paratopic vehicles, that are extremely appealing for enhanced specificities and cell surface avidity in receptor-mediated drug delivery [38-42]. In vehicles with such combined functional agents, the residual CPP activities of polyarginines might enhance penetrability and endosomal escape more efficiently than completely unspecific fusogenic peptides that such as HA2, tend to impair selectivity of the accompanying cell ligands [30].

Acknowledgments: Protein production has been partially performed by the ICTS "NANBIOSIS", more specifically by the Protein Production Platform of CIBER-BBN/ IBB, at the UAB SepBioES scientific-technical service (http://www.nanbiosis.es/unit/u1-protein-production-platform-ppp/). Nanoparticle size determination was performed at the NANBIOSIS Biomaterial Processing and Nanostructuring Unit of CIBER-BBN (http://www.nanbiosis.es/portfolio/u6-biomaterial-processing-and-nanostructuring-unit).

Legends

Figure 1. Production and preliminary characterization of Rn-containing GFP protein versions. A. Generic scheme of the modular protein construction. The linker sequence is GGGNS, and the precise amino acid sequence of the Rn peptides is indicated in the box. B Coomassie blue-stained PAGE-SDS gel showing the integrity and purity of Rn-GFP-H6 recombinant proteins upon affinity chromatography (left). Proteolytic stability was confirmed by Western blot (right). The type of buffer used is indicated at the top, as well as the number of amino-terminal arginine residues in each protein. The molecular mass of relevant markers (M) is indicated in kDa. C. Fluorescence emission spectra of equal amounts of each protein.

Figure 2. Assembling of Rn-containing GFP proteins. DLS size determination of protein self-assembling. The size of unassembled proteins (6-8 nm) might correspond to GFP dimers acting as building blocks. In the inset, peak (expressed in nm) and PDI as mean values and standard deviation ($\overline{x}$ ± SD) for each protein sample obtained by DLS.

Figure 3. Ultrastructural analysis of Rn-containing GFP proteins. FESEM and TEM imaging of protein nanoparticles formed in *Dextrose* buffer. Representative fields of R7-GFP-H6 samples in *NaCl* buffer are also shown as negative controls. Bars size is 20 nm. The size of the nanomaterials (in nm) determined from images is indicated as mean values and standard deviation ($\overline{x}$ ± SD) for each protein sample and EM technique (nd indicates that the material has not been detected).

Figure 4. Cell penetrability of Rn-containing GFP proteins. A. Intracellular fluorescence accumulated in $CXCR4^+$ HeLa cells upon exposure to unassembled and assembled Rn-based proteins, for different times and protein doses. Symbols are ** $p < 0.01$; * $p < 0.05$. B. Conventional confocal images of target HeLa cells exposed for 24 h to 2 M of either R7 or R9-based proteins. The green signal results from the protein fluorescence, while red signals label membranes and blue signals the cell nuclei. Arrows indicate yellow merging signals. C. 3D reconstructions based on stacks of 20–30 sections of protein-exposed HeLa cells. In the insets, orthogonal sections of 3D confocal images Bars indicate 5 m.

Figure 5. CXCR4-dependence in the internalization of Rn-based protein materials. A. Accumulation of intracellular fluorescence into $CXCR4^+$ HeLa cells associated to different Rn-based materials, and AMD3100-mediated inhibition of the process. The signals were recorded at two different times after exposure. Symbols are ** $p < 0.01$; * $p < 0.05$. B. DLS size measurements of R7 and R9-based proteins upon incubation with SDS for 10 min.

Figure 6. Internalization of Rn-based protein materials. A. Intracellular florescence in HeLa cells and percentage of AMD3100-mediated uptake inhibition upon 2 h and 24 h of exposure to Rn-based nanoparticles. B. Plotting of main kinetic parameters of protein uptake, namely L, τ, , k and v (see Eq 1-3) versus the number of arginine residues at the N-terminus of proteins (Rn). Uptake experiments in presence of AMD3100 are indicated (+AMD).

References

[1] Chan DC, Kim PS. HIV entry and its inhibition. Cell. 1998;93:681-4.

[2] Burger JA, Kipps TJ. CXCR4: a key receptor in the crosstalk between tumor cells and their microenvironment. Blood. 2006;107:1761-7.

[3] Unzueta U, Cespedes MV, Vazquez E, Ferrer-Miralles N, Mangues R, Villaverde A. Towards protein-based viral mimetics for cancer therapies. Trends in biotechnology. 2015;33:253-8.

[4] Barbieri F, Bajetto A, Florio T. Role of chemokine network in the development and progression of ovarian cancer: a potential novel pharmacological target. Journal of oncology. 2010;2010:426956.

[5] Choi WT, Yang Y, Xu Y, An J. Targeting chemokine receptor CXCR4 for treatment of HIV-1 infection, tumor progression, and metastasis. Current topics in medicinal chemistry. 2014;14:1574-89.

[6] Saccardo P, Villaverde A, Gonzalez-Montalban N. Peptide-mediated DNA condensation for non-viral gene therapy. Biotechnology advances. 2009;27:432-8.

[7] Tanaka G, Nakase I, Fukuda Y, Masuda R, Oishi S, Shimura K, et al. CXCR4 stimulates macropinocytosis: implications for cellular uptake of arginine-rich cell-penetrating peptides and HIV. Chemistry & biology. 2012;19:1437-46.

[8] Unzueta U, Seras-Franzoso J, Cespedes MV, Saccardo P, Cortes F, Rueda F, et al. Engineering tumor cell targeting in nanoscale amyloidal materials. Nanotechnology. 2017;28:015102.

[9] Unzueta U, Cespedes MV, Ferrer-Miralles N, Casanova I, Cedano J, Corchero JL, et al. Intracellular CXCR4(+) cell targeting with T22-empowered protein-only nanoparticles. International journal of nanomedicine. 2012;7:4533-44.

[10] Rueda F, Cespedes MV, Conchillo-Sole O, Sanchez-Chardi A, Seras-Franzoso J, Cubarsi R, et al. Bottom-Up Instructive Quality Control in the Biofabrication of Smart Protein Materials. Advanced materials. 2015;27:7816-22.

[11] Unzueta U, Ferrer-Miralles N, Cedano J, Zikung X, Pesarrodona M, Saccardo P, et al. Non-amyloidogenic peptide tags for the regulatable self-assembling of protein-only nanoparticles. Biomaterials. 2012;33:8714-22.

[12] Serna N, Cespedes MV, Saccardo P, Xu Z, Unzueta U, Alamo P, et al. Rational engineering of single-chain polypeptides into protein-only, BBB-targeted nanoparticles. Nanomedicine : nanotechnology, biology, and medicine. 2016;12:1241-51.

[13] Serna NC, M; Sánchez-García, L; Unzueta, U; Sala, R; Sánchez-Chardi, A; Cortés, F; Ferrer-Miralles, N; Mangues, R; Vázquez, E; Villaverde, A. Peptide-Based Nanostructured Materials with Intrinsic Proapoptotic Activities in CXCR4+ Solid Tumors. Advanced Functional Materials. 2017;27:1700919.

[14] Serna N, Sanchez-Garcia L, Sanchez-Chardi A, Unzueta U, Roldan M, Mangues R, et al. Protein-only, antimicrobial peptide-containing recombinant nanoparticles with inherent built-in antibacterial activity. Acta biomaterialia. 2017. 60:256-63.

[15] Vazquez E, Roldan M, Diez-Gil C, Unzueta U, Domingo-Espin J, Cedano J, et al. Protein nanodisk assembling and intracellular trafficking powered by an arginine-rich (R9) peptide. Nanomedicine. 2010;5:259-68.

[16] Kim HY, Hwang JY, Kim SW, Lee HJ, Yun HJ, Kim S, et al. The CXCR4 Antagonist AMD3100 Has Dual Effects on Survival and Proliferation of Myeloma Cells In Vitro. Cancer research and treatment : official journal of Korean Cancer Association. 2010;42:225-34.

[17] Pesarrodona M, Crosas E, Cubarsi R, Sanchez-Chardi A, Saccardo P, Unzueta U, et al. Intrinsic functional and architectonic heterogeneity of tumor-targeted protein nanoparticles. Nanoscale. 2017;9:6427-35.

[18] Pesarrodona M, Fernandez Y, Foradada L, Sanchez-Chardi A, Conchillo-Sole O, Unzueta U, et al. Conformational and functional variants of CD44-targeted protein nanoparticles bio-produced in bacteria. Biofabrication. 2016;8:025001.

[19] Futaki S, Nakase I. Cell-Surface Interactions on Arginine-Rich Cell-Penetrating Peptides Allow for Multiplex Modes of Internalization. Accounts of chemical research. 2017;50:2449-56.

[20] Nakase I, Niwa M, Takeuchi T, Sonomura K, Kawabata N, Koike Y, et al. Cellular uptake of arginine-rich peptides: roles for macropinocytosis and actin rearrangement. Molecular therapy : the journal of the American Society of Gene Therapy. 2004;10:1011-22.

[21] Jung YH, Lee DY, Cha W, Kim BH, Sung MW, Kim KH, et al. Antitumor effect of CXCR4 antagonist AMD3100 on the tumorigenic cell line of BHP10-3 papillary thyroid cancer cells. Head & neck. 2016;38:1479-86.

[22] Atif SM, Hasan I, Ahmad N, Khan U, Owais M. Fusogenic potential of sperm membrane lipids: nature's wisdom to accomplish targeted gene delivery. FEBS letters. 2006;580:2183-90.

[23] Ahmad N, Masood AK, Owais M. Fusogenic potential of prokaryotic membrane lipids. Implication in vaccine development. European journal of biochemistry. 2001;268:5667-75.

[24] Ye J, Liu E, Yu Z, Pei X, Chen S, Zhang P, et al. CPP-Assisted Intracellular Drug Delivery, What Is Next? International journal of molecular sciences. 2016;17.

[25] Ye J, Shin MC, Liang Q, He H, Yang VC. 15 years of ATTEMPTS: a macromolecular drug delivery system based on the CPP-mediated intracellular drug delivery and antibody targeting. Journal of controlled release : official journal of the Controlled Release Society. 2015;205:58-69.

[26] Dai Q, Bertleff-Zieschang N, Braunger JA, Bjornmalm M, Cortez-Jugo C, Caruso F. Particle Targeting in Complex Biological Media. Advanced healthcare materials. 20118; 7.

[27] Mangues RV, E; Villaverde, A. Targeting in Cancer Therapies. Medical Sciences. 2016;4:6.

[28] Vazquez E, Ferrer-Miralles N, Villaverde A. Peptide-assisted traffic engineering for nonviral gene therapy. Drug discovery today. 2008;13:1067-74.

[29] Ferrer-Miralles N, Vazquez E, Villaverde A. Membrane-active peptides for non-viral gene therapy: making the safest easier. Trends in biotechnology. 2008;26:267-75.

[30] Sanchez-Garcia L, Serna N, Mattanovich M, Cazzanelli P, Sanchez-Chardi A, Conchillo-Sole O, et al. The fusogenic peptide HA2 impairs selectivity of CXCR4-targeted protein nanoparticles. Chemical communications. 2017;53:4565-8.

[31] Guidotti G, Brambilla L, Rossi D. Cell-Penetrating Peptides: From Basic Research to Clinics. Trends in pharmacological sciences. 2017;38:406-24.

[32] Niu Z, Tedesco E, Benetti F, Mabondzo A, Montagner IM, Marigo I, et al. Rational design of polyarginine nanocapsules intended to help peptides overcoming intestinal barriers. Journal of controlled release : official journal of the Controlled Release Society. 2017;263:4-17.

[33] Kawaguchi Y, Takeuchi T, Kuwata K, Chiba J, Hatanaka Y, Nakase I, et al. Syndecan-4 Is a Receptor for Clathrin-Mediated Endocytosis of Arginine-Rich Cell-Penetrating Peptides. Bioconjugate chemistry. 2016;27:1119-30.

[34] Cespedes MV, Unzueta U, Alamo P, Gallardo A, Sala R, Casanova I, et al. Cancer-specific uptake of a liganded protein nanocarrier targeting aggressive CXCR4+ colorectal cancer models. Nanomedicine : nanotechnology, biology, and medicine. 2016;12:1987-96.

[35] de Nigris F, Crudele V, Giovane A, Casamassimi A, Giordano A, Garban HJ, et al. CXCR4/YY1 inhibition impairs VEGF network and angiogenesis during malignancy. Proceedings of the National Academy of Sciences of the United States of America. 2010;107:14484-9.

[36] Murakami T, Cardones AR, Hwang ST. Chemokine receptors and melanoma metastasis. Journal of dermatological science. 2004;36:71-8.

[37] Song JS, Kang CM, Kang HH, Yoon HK, Kim YK, Kim KH, et al. Inhibitory effect of CXC chemokine receptor 4 antagonist AMD3100 on bleomycin induced murine pulmonary fibrosis. Experimental & molecular medicine. 2010;42:465-72.

[38] Kontermann RE. Dual targeting strategies with bispecific antibodies. mAbs. 2012;4:182-97.

[39] Kontermann RE, Brinkmann U. Bispecific antibodies. Drug discovery today. 2015;20:838-47.

[40] Muller D, Kontermann RE. Recombinant bispecific antibodies for cellular cancer immunotherapy. Current opinion in molecular therapeutics. 2007;9:319-26.

[41] Weidle UH, Kontermann RE, Brinkmann U. Tumor-antigen-binding bispecific antibodies for cancer treatment. Seminars in oncology. 2014;41:653-60.

[42] Unzueta U, Serna N, Sanchez-Garcia L, Roldan M, Sanchez-Chardi A, Mangues R, et al. Engineering multifunctional protein nanoparticles by in vitro disassembling and reassembling of heterologous building blocks. Nanotechnology. 2017;28:505102.

Figure 1

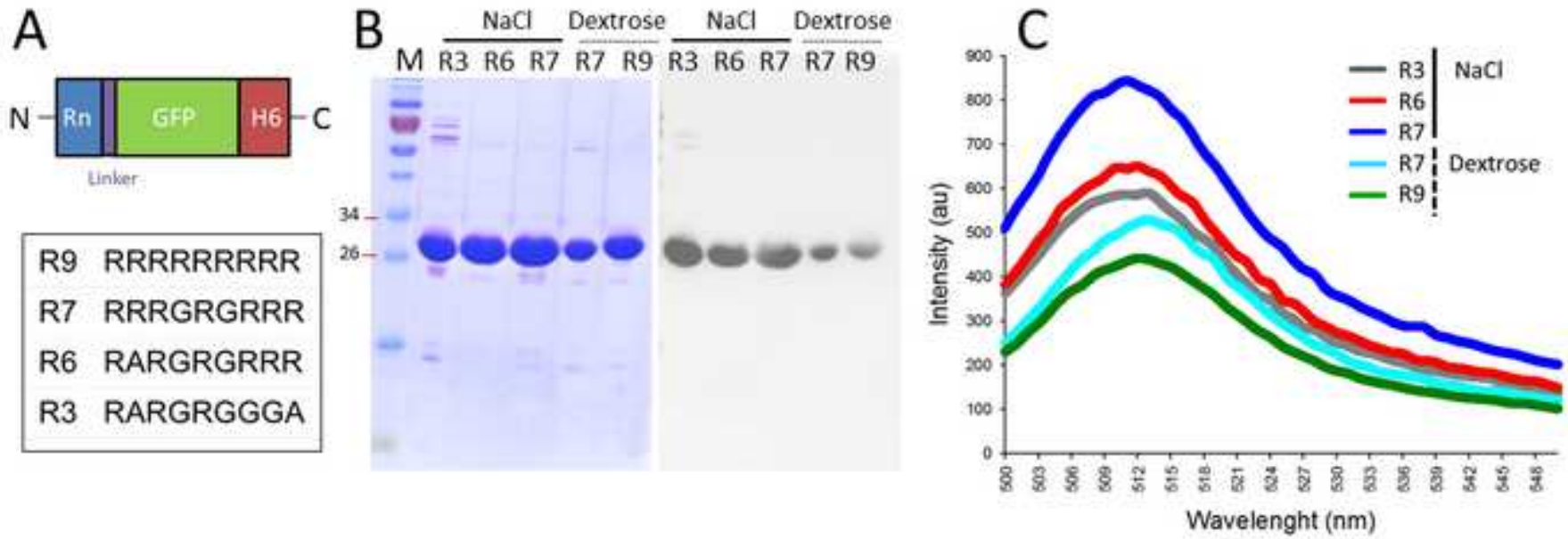
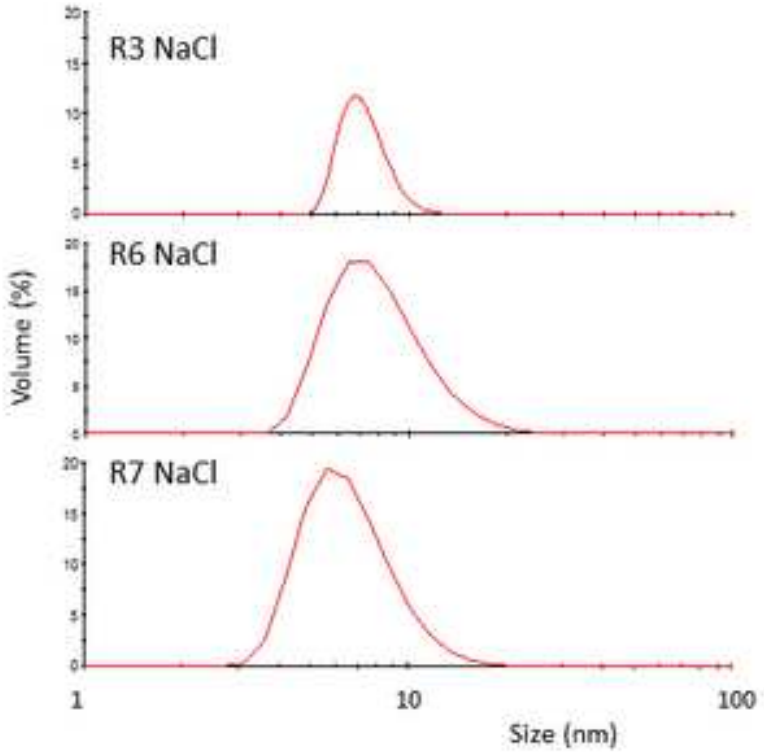
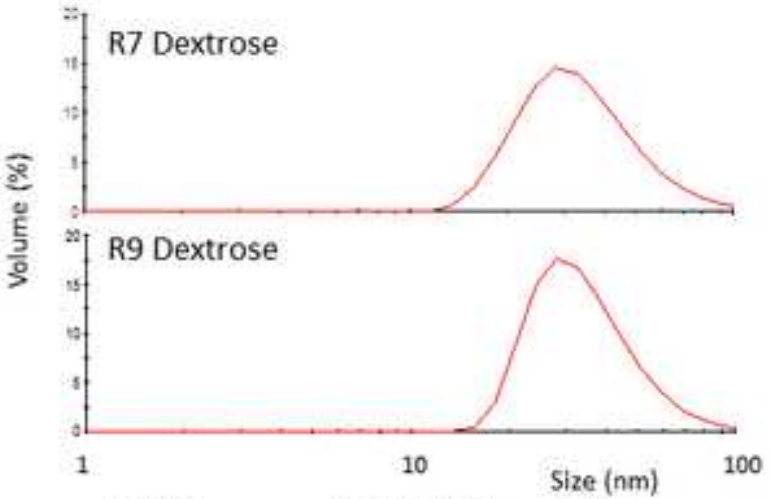

A
N
Rn
GFP
H6
C
Linker
R9 RRRRRRRRR
R7 RRRGRGRRR
R6 RARGRGRRR
R3 RARGRGGGA
B
NaCl
Dextrose
NaCl
Dextrose
M R3 R6 R7 R7 R9 R3 R6 R7 R7 R9
34
26
C
R3
R6
R7
R7
R9
NaCl
Dextrose
Intensity (au)
Wavelenght (nm)

Figure 2





Protein	Peak size (nm)	pDi
R3 NaCl	7.27 ± 0.23	0.67 ± 0.02
R6 NaCl	8.22 ± 0.09	0.34 ± 0.03
R7 NaCl	6.70 ± 0.16	0.67 ± 0.05
R7 Dextrose	27.64 ± 13.54	0.45 ± 0.11
R9 Dextrose	34.90 ± 0.23	0.15 ± 0.01

Figure 3

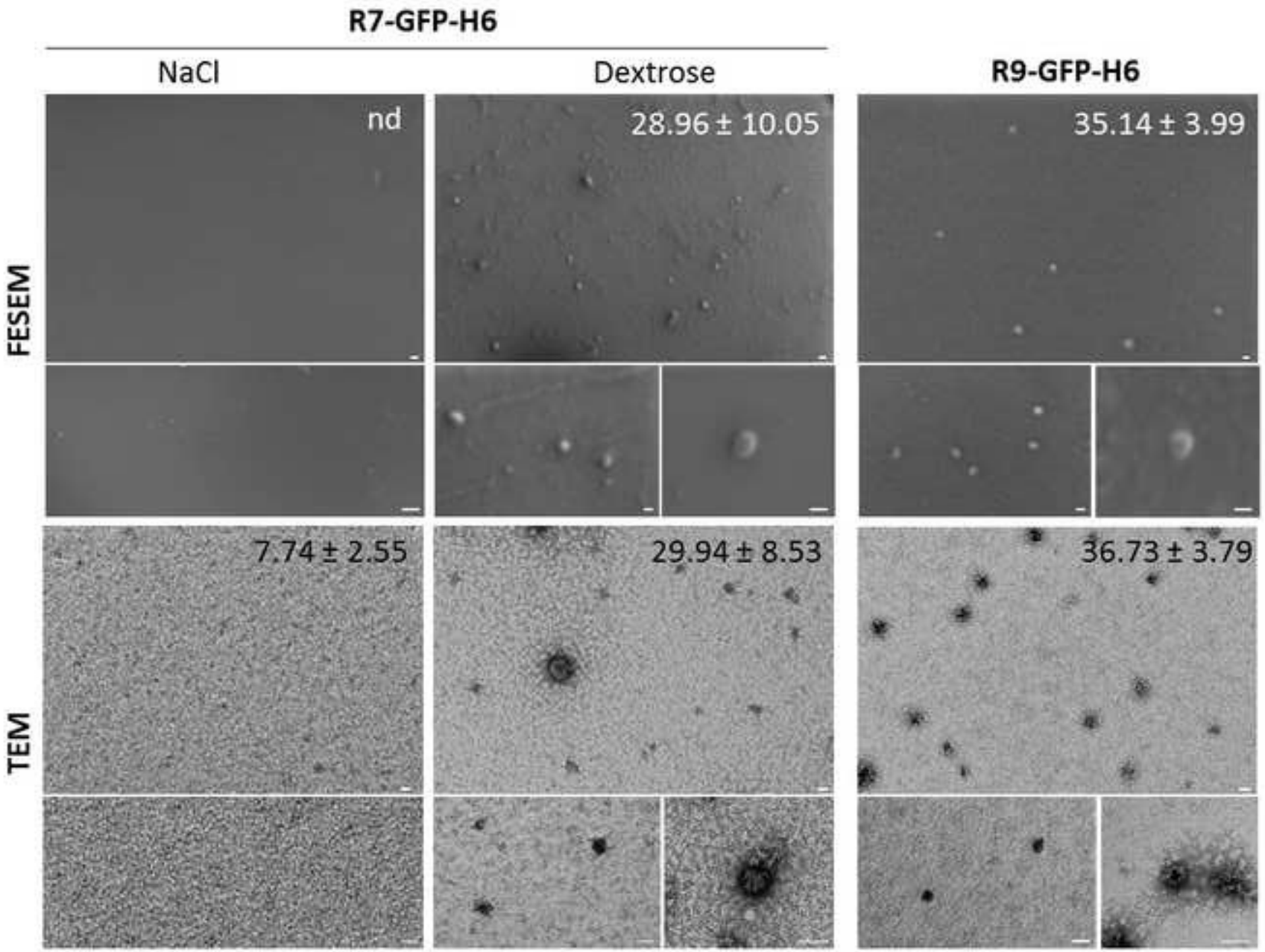

R7-GFP-H6
NaCl
Dextrose
R9-GFP-H6
FESEM
TEM
nd
28.96 ± 10.05
35.14 ± 3.99
7.74 ± 2.55
29.94 ± 8.53
36.73 ± 3.79

Figure 4

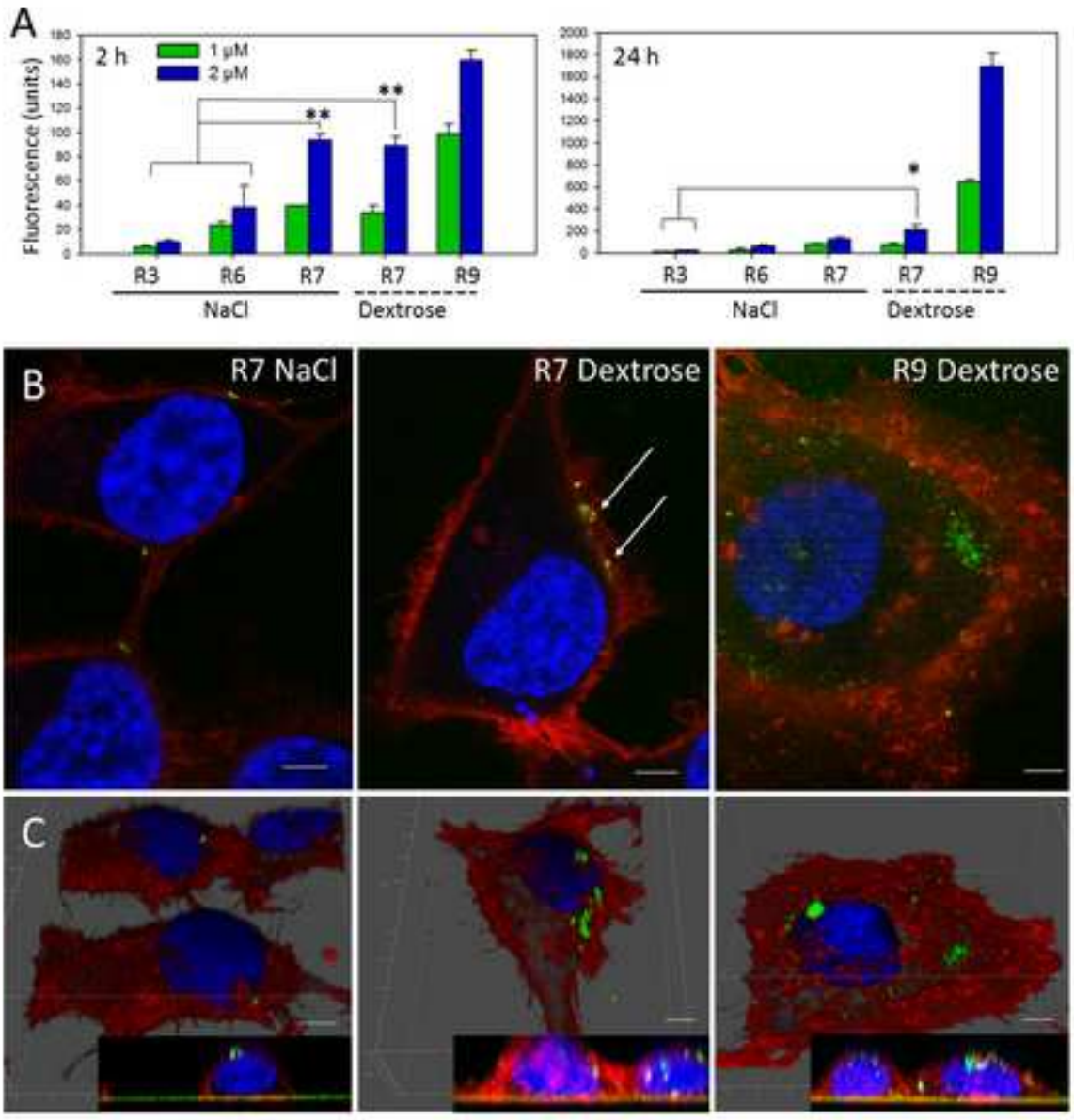

A
2 h
1 µM
2 µM
Fluorescence (units)
**
**
R3
R6
R7
R7
R9
NaCl
Dextrose
24 h
*
R3
R6
R7
R7
R9
NaCl
Dextrose
B
R7 NaCl
R7 Dextrose
R9 Dextrose
C

Figure 5

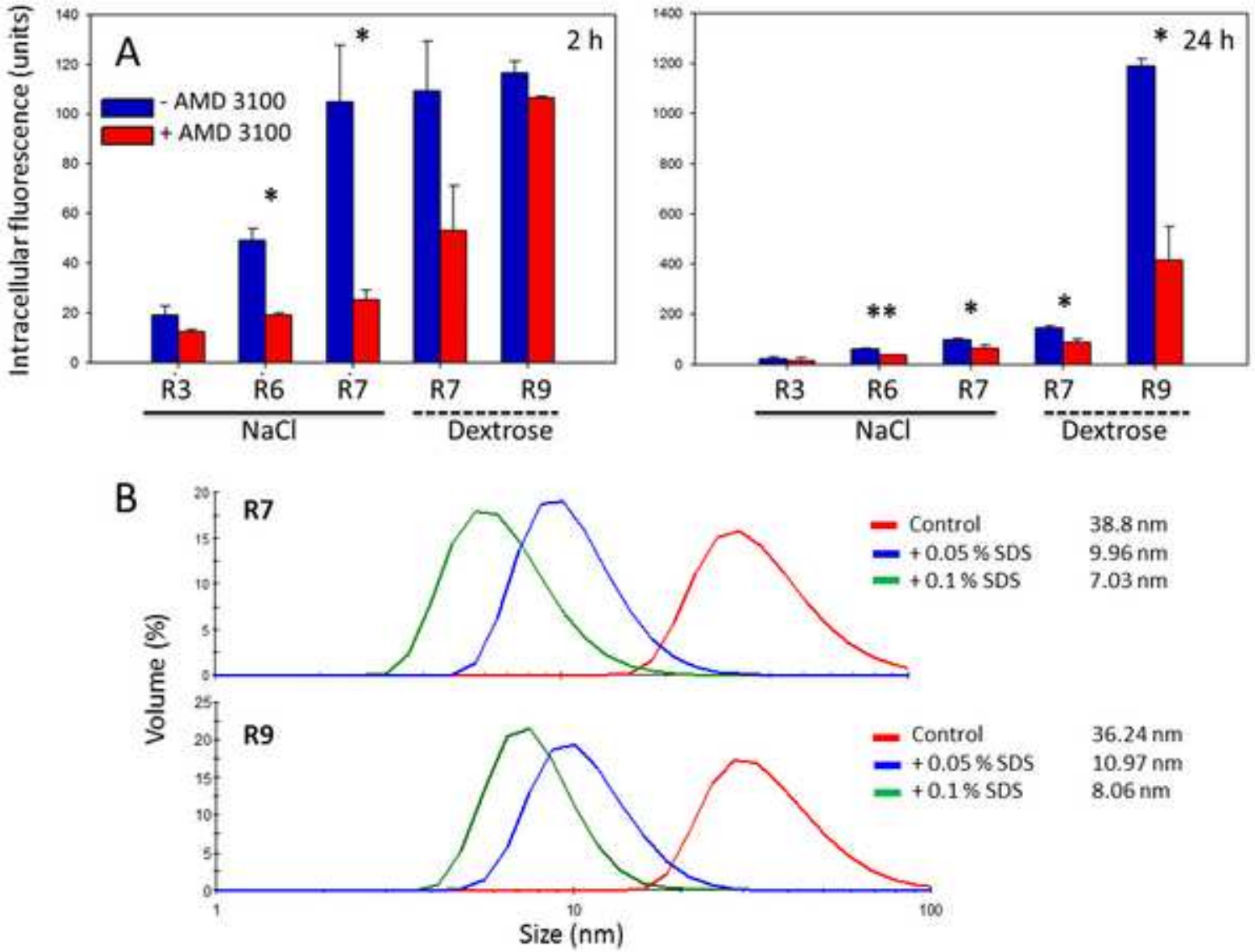
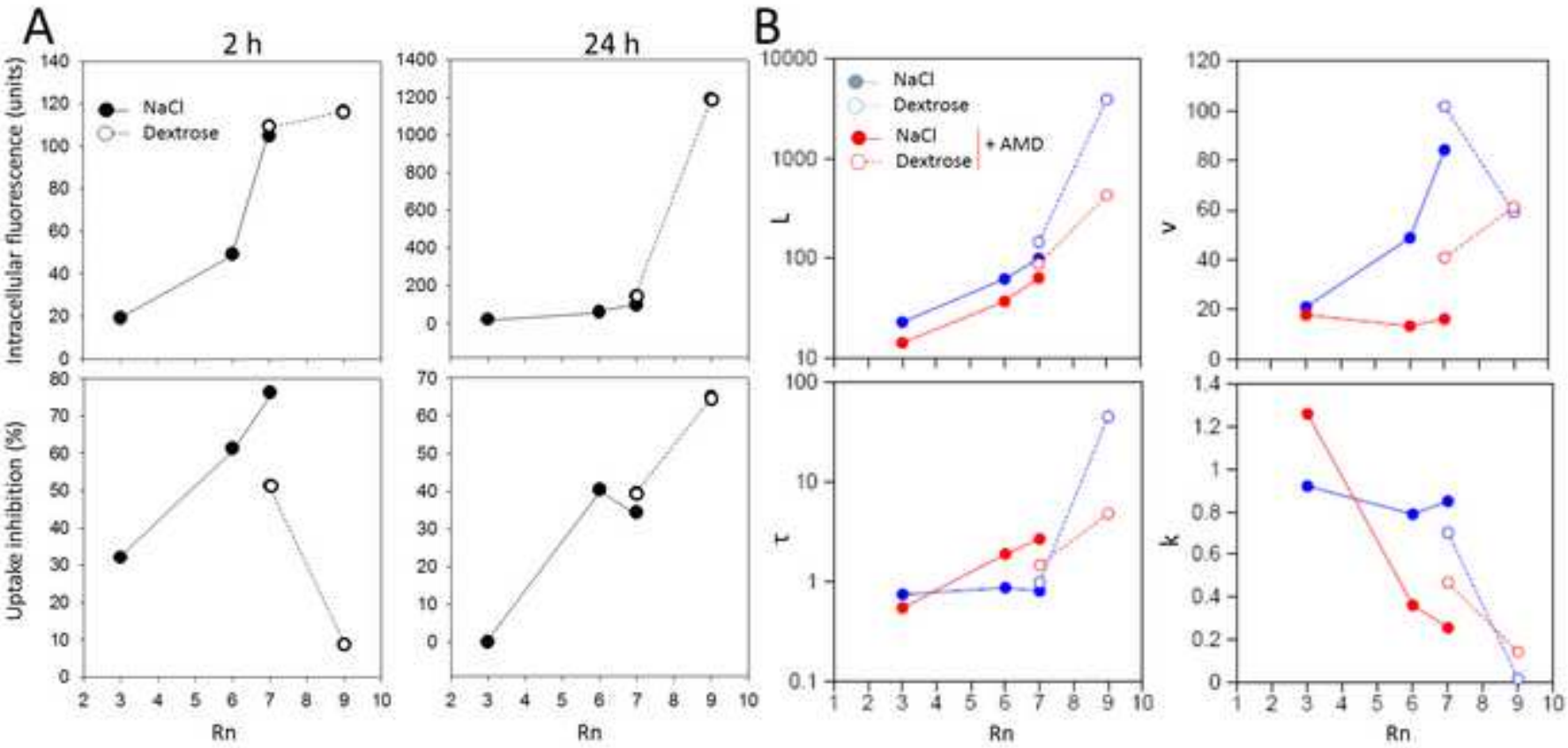

A
2 h
24 h
- AMD 3100
+ AMD 3100
Intracellular fluorescence (units)
R3
R6
R7
R7
R9
NaCl
Dextrose
B
R7
R9
Control
+ 0.05 % SDS
+ 0.1 % SDS
38.8 nm
9.96 nm
7.03 nm
36.24 nm
10.97 nm
8.06 nm
Volume (%)
Size (nm)

Figure 6


A
2 h
24 h
Intracellular fluorescence (units)
Uptake inhibition (%)
NaCl
Dextrose
Rn
B
NaCl
Dextrose
NaCl
Dextrose
+ AMD
L
v
τ
k
Rn